Solar Power Forecasting Performance – Towards Industry Standards

V. Kostylev and A. Pavlovski

Abstract-- **Due to the rapid increase in deployment and high penetration of solar power generation worldwide, solar power generation forecasting has become critical to variable generation integration planning, and within utility and independent system operator (ISO) operations. Utilities and ISOs require day ahead and hour ahead as well as intra-hour solar power forecasts for core operations - solar power producers and energy traders also require high quality solar power forecasts.**

As a result of the erroneously perceived simplicity of solar radiation forecasting, very often non-repeatable, poorly explained or obscure estimates of solar power forecast performance are used. This creates uncertainty with the quality of forecasting service, as well as unrealistic expectations of possible forecast precision. As a result, there is an immediate need for defining a common methodology for evaluating forecast performance, establishing verification procedures, and setting common standards for industry-approved quality of solar forecast performance.

Solar power forecast quality claims can be easily verified when the source of forecast is known. Most often the offered power generation forecasts are based on publically available results of Numerical Weather Prediction (NWP) models and on the use of empirical relationships between solar resource and generated power at a specific plant. The quality of these forecasts is limited by the quality of the NWP models utilized, which is known. Less frequently, solar radiation is estimated based on proprietary models such as satellite-based or total sky imager–based cloud cover and radiation forecasts. In such cases, there are also known limits to the accuracy of prediction which can help objectively evaluate claims of the forecast service companies.

This paper is proposing a set of standards for evaluating intra hour, hour ahead, day ahead and week ahead solar power forecast performance. The proposed standards are based on sound methodologies and extensive field practice and offer a solid ground for reliable inter-agency comparisons of forecast performance.

Index Terms **— Forecast standards, RMSE, bias, persistence, NWP, performance evaluation.**

I. Introduction

DUE to the rapid increase in deployment and high penetration of solar power in electricity grids worldwide, solar power generation forecasting has become critical to variable generation integration planning and within utility and ISO operations. State-of-the-art solar power forecasting is seen as a major tool to address the risks related to the high share of variable solar generation within the electricity mix, limiting curtailment of generation and reducing idle backup capacity. Moreover, it is seen as an important component of the Smart Grid toolbox enabling efficient demand-side management and demand response measures.

Currently all major participants in the electricity value chain: solar power producers, utilities and ISOs – show considerable interest in high-precision solar power forecasts and imply their own technical requirements to forecast outputs and quality. Utilities and ISOs require solar power forecasts for core operations; solar power producers, in certain jurisdictions, are mandated to provide power forecasts by their power purchase agreements.

There are, however, no industry standards defining applicable and credible data sources, forecast technologies and validating procedures established thus far. This creates uncertainty with the quality of forecasting service, as well as unrealistic expectations of possible forecast precision. To enable the industry with high-precision solar forecasts there is an immediate need for defining a common methodology for evaluating forecast performance, establishing verification procedures and setting common standards for industry-approved quality of solar forecast performance.

At present, there are no widely recognized standards or recommended practices and procedures. There is an immediate need in developing recommended practices for solar radiation forecasting techniques which would assist in producing more reliable forecasts, optimally utilizing the current levels of meteorological science and computer technologies. Such recommendations depend on the degree of understanding of limitation of the existing approaches, particularly NWP and statistical methods commonly used by the solar forecast industry. This publication presents an approach to developing industry standards in solar power forecasting with the focus on evaluation of solar forecast performance.

A. *Solar Power Forecast Time Scales*

The most critical technical requirements to solar power forecasting are defined by forecast time scales demanded by the electricity value chain participants. This group, however, does not provide clear and detailed explanation of the practical uses of these forecasts in reference to practical

V. Kostylev and A. Pavlovski are with Green Power Labs Inc., One Research Dr. Dartmouth, Nova Scotia, Canada (e-mail: vkostylev@greenpowerlabs.com).

benefits.

Most common industry-requested operational forecasts and their corresponding granularity are the following:

- Intra–Hour: 15 minutes to 2 hours ahead with 30 seconds to 5 minute granularity (relates to ramping events, variability related to operations)
- Hour Ahead: One to 6 hours ahead with hourly granularity (related to load following forecasting)
- Day Ahead: One to 3 days ahead with hourly granularity (relates to unit commitment, transmission scheduling, and day ahead markets)
- Medium-term: Week to 2 months ahead, with daily granularity (hedging, planning, asset optimisation)
- Long-term: typically one or more years, with diurnal monthly and annual granularity (long-term time series analysis, resource assessment, site selection, and bankable documentation)

Achieving high-accuracy forecasts at each of these time scales imposes specific requirements to applicable solar radiation models, data sources, and forecasting techniques converting available data into quality solar power forecasts. Day Ahead forecasts are needed for operational planning, switching sources, programming backup, and short-term power purchases, as well as for planning of reserve usage, and peak load matching [1]. Medium-term forecasts are concerned with planning, plant optimization, risk assessment, while Long-term forecasts (also known as resource assessment) are targeting return on investment estimates. Along this time continuum different forecasting approaches and evaluation techniques should be used.

II. Evaluation of forecast quality

A. *Interpretation of statistical measures*

Forecast performance testing is often approached from established analytical practices without clear understanding of relevance of the performance measures to particular management applications (e.g. plant operation and management, energy trading etc.). This leads to reporting a variety of metrics for a subset of forecast scales of practical importance. Mean absolute error (MAE), mean bias error (MBE) and root mean square error (RMSE) are the three most commonly used statistics to test solar radiation forecast performance. In evaluations of solar radiation models these are used together, interchangeably or added to each other for evaluation of total score (e.g. [2]) and expressed either in absolute values (e.g. W/m2) or as fraction of some metrics of observed data. Percent departure from a persistence-based model or climatological means further complicates interpretation of forecast quality. Additionally, of a particular significance to the power industry, are forecasts evaluated at a certain time ahead, e.g. “hour ahead” and “day ahead” which should be clearly differentiated from “hourly” and “daily” forecasts.

In-depth review on the use of various statistics in forecast evaluation is a matter of a separate study (being prepared for publication elsewhere) and we will offer only several observations on metrics deserving the highest attention. RMSE statistic is the most commonly reported in forecast accuracy claims. While it is a good measure of forecast uncertainty the published values are often non-informative because: 1) When expressed as relative RMSE (%) details are often missing on normalization (to mean or range of observed values for the analysis period), 2) Often decrease in RMSE compared to other methods or persistence is discussed instead of actual values, and 3) In analyses of power generation RMSE is often expressed in percent of rated power rather than observed power output.

MAE has a specific meaning and it is loosely related to RMSE because it puts less emphasis on the extreme discrepancies between forecasted and observed values. MBE also has different meaning and value to forecast performance evaluation. It relates more to general over, or under-prediction over the analysis time span, rather than to predictive power of forecast. Therefore, the values of these three statistics tell different stories about forecast quality and their published values depend on the time span of validation analysis, use of sunlight hours only, and data aggregation techniques which could render these values incomparable between studies.

B. *Use of Naïve Models*

A number of naïve approaches is commonly used as benchmarks for demonstrating relative accuracy of forecasts. Persistence, as a forecasting principle, can be used for producing forecasts for minutes to hours-ahead time scales, and less commonly to forecast day ahead at clear sky conditions. At the same time, persistence–based forecasts are the most commonly used for comparison with commercial forecast performance. Similarly to that of other performance metrics, the methodological details of ‘persistence–based comparisons’ are commonly poorly described.

General logic of the persistence based testing is the assumption of no change, i.e. forecasted conditions assumed similar to current. Various variations of persistence-based modeling in solar radiation forecasts are 1) assuming that the forecasted values equal current values (e.g. [3]) 2) assuming that the relative values of forecasted values equal current and adjusting them by the stage of diurnal cycle 3) using recent trends of change (e.g. in ARIMA approaches) or 4) using climatology as a benchmark, thereby assuming long-term persistence of climatological means. To further complicate things, the persistence principles can be applied to solar photovoltaic (PV) production, solar radiation, cloud cover [4] or the whole ensemble of weather prediction variables. The source of information used for persistence modeling also affects the results – solar radiation or cloud cover measurements obtained at the site provide better persistence estimates then e.g. NWP-derived values, consequently affecting the Forecast/Persistence performance ratios. Persistence-based forecasts derived from ground measured clear sky index (ratio of ground based global and clear sky radiation) and satellite-based cloud index produce dramatically different results. Satellite-based persistence model outperforms persistence based on ground data due to the high resolution of spatial information provided by satellite images [3, 5]. Sufficiently advanced "persistence" model however may become impractical as a benchmark over which an improvement is expected, while overly naïve approaches based on modeled data could be too easy to outcompete. Based on this, we suggest that the strictest benchmark for persistence–based forecast evaluation would be provided by a) using recent empirical measurements of solar radiation as inputs, and producing persistence based forecast assuming that the relative values of forecasted values vary with the stage of diurnal solar cycle (with zenith angle). This approach provides sufficient simplicity of implementation while preserving realism of input values.

III. Effects of Forecast Models on Accuracy

1) NWP-based forecasting

Numerical weather prediction models such as Global Environmental Multiscale Model (GEM), Global Forecast System (GFS), North American Mesoscale Model (NAM), European Centre for Medium Range Forecasts (ECMWF) model, etc. are commonly used to derive solar radiation forecasts. These global models provide forecast for a number of variables useful for modeling solar radiation at a 6 hour refresh rate and at 1 to 3 hours granularity, with forecast horizons from 48 to 180 hours or longer. Spatial resolution of these models is variable and on average each forecast grid cell is on a scale of hundreds of square kilometers. For example, ECMWF delivers forecast up to 10 days ahead, but in a maximum spatial resolution of approximately 60 km x 60 km. The radiative transfer models in most NWPs are producing output only hourly (NAM) or every 3 (GFS and ECMWF) [6] or 4 hours (in High Resolution Rapid Refresh (HRRR) model). These outputs are also not tuned to a particular forecast location but apply to the large grid cells. Cloud cover information derived from NWPs is commonly used for integration with clear sky model [7] to produce location specific forecast. Even with HRRR, which is delivered at 3x3km resolution, hourly updated clouds are interpolated to this fine resolution from predictions carried at 30x30 km grid.

Performance of NWP as a source of data for solar radiation forecast is highly variable, and across several models, can vary by 200% (Fig. 1) (e.g. from RMSE 20% to 60% for the day-ahead forecast) when applied to different locations.

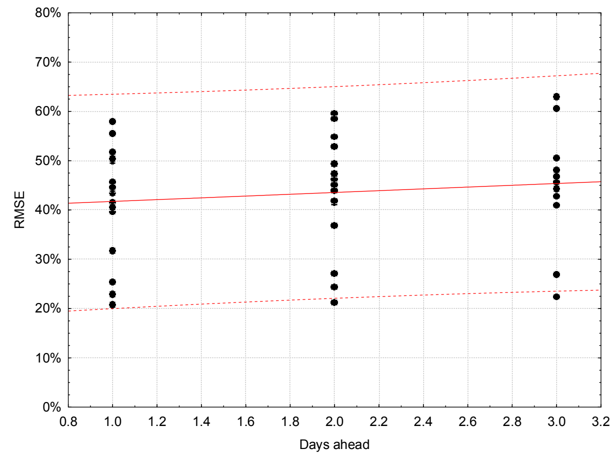

Fig. 1. RMSE of 6 different NWP approaches and persistence at 4 geographic areas, normalized to the mean ground irradiance (based on data from [4]).

Because of the nature of spatio-temporal coupling, use of NWP is optimal at forecast time scales from 6 hours and longer. Lave and Kleissl [8] present spatio-temporal coherence spectra which show that sites 60 km or more apart were uncorrelated on timescales shorter than 12 hours and sites that were only 19 km apart were uncorrelated on timescales shorter than 3 hours. Perez et al. [9] show that the distance at which solar radiation at station pairs become uncorrelated is almost a linear function of the considered time scale with distances of 1 km corresponding to 1 minute and 10 km to 15 minutes. Despite large variability in reported absolute values of decorrelation distance vs. time, the coarser resolution NWPs, such as GEM, could be used for producing forecasts for further ahead times than high resolution NWPs, e.g. HRRR, and higher spatial resolution methods should be applied to achieve better accuracy at finer temporal scales.

2) Satellite-based forecasting

Cloudiness is the most important factor affecting the total amount of solar radiation reaching Earth's surface and in determining intermittency or solar power production. Half-hourly visible spectrum images of cloud cover provide a convenient way to evaluate the amount of solar radiation incident of the ground. These images provide near 1 square kilometer resolution for cloud cover at satellite nadir, and the methodology was successfully used to produce climatological NSRDB-SUNY dataset for US (e.g. [10]). Typical RMSEs of the satellite-derived surface irradiance compared to ground observations are 20-25 % for hourly data, while daily and monthly values generally show uncertainties of 8-12 % and 5-7 %, respectively [5]. Monthly averaged satellite-derived GHI shows MBE of 0.84% and RMSE of 5.25% when compared to ground-based measurements [11].

In forecasting applications, motion vectors are used to interpolate future position of clouds over ground which can provide accurate forecasts up to 6 h ahead [12] with temporal granularity of minutes. The usefulness of this approach for explaining radiation variability is the highest under partial cloudiness, when abrupt changes may trigger an attenuation of 60–70% in GHI and ≈100% in DNI [13] or under constant clouds. From 1 to 6 hours ahead satellite-based models produce results superior to NWP-based and persistence based models [12].

Limitations of this approach are in infrequent updates of the original images, bad geographic registration of satellite images, poor understanding of cloud altitudes which pose particular problems for sunrise and sunset predictions, and a set of challenges posed by estimating clearness index through calculation of dynamic pixel range. High bare ground albedo is the most common problem weakening the approach in arid environments and as a result of seasonal snow cover. Unrealistic assumption of steady state cloud cover (i.e. that cloud pattern moves as a single, unchangeable layer) is probably the main weakness of the approach. Clouds moving at different altitudes and their shadows on the ground as well as on top of other clouds produce an additional challenge in correct interpolation of the next cloud image and determination of clearness index.

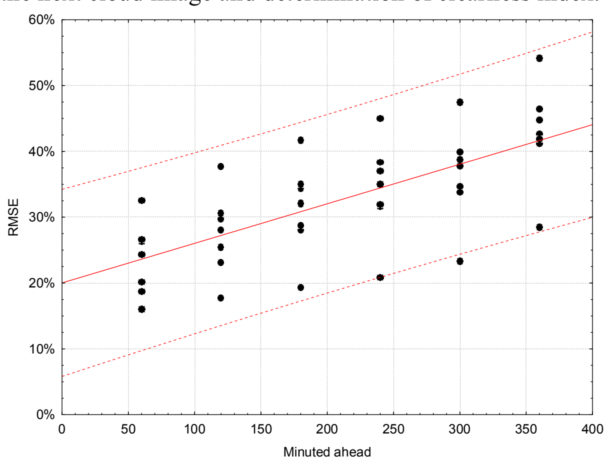

Fig. 2. Satellite-based GHI forecast RMSE as a function of time ahead, based on data in [12] for 7 locations. While original data presents w/m2, we have calculated RMSE as % of mean observed. 0.95 prediction intervals are shown.

3) *Total sky imager*

Similarly to satellite-based forecast approaches the use of total sky imagers provides near real-time detection of clouds and capacity to predict cloud movement/position in the near future. Presence of clouds at a particular pixel of hemispherical sky image is identified using various thresholding algorithms [14][15]. Cloud motion vectors are generated by cross-correlating consecutive sky images and used to predict cloud locations short time ahead, dependent on velocity of cloud movement. This approach allows for high spatial and temporal resolution in GHI forecasts at timescales shorter than about 5-min allowing discrimination between sites separated by few km [8].

Chow et al. [16] found a 50–60% reduction in forecast error in TSI based forecast compared to persistence for 30 seconds ahead forecast. Cloud forecast error increased with increasing forecast horizon due to high cloud cover variability over the coastal site. Mean total matching error (between forecasted and observed cloud maps) was reported in the range 6% - 30% for 30 seconds to 5 minutes ahead correspondingly [16]. It is not clear how exactly this error translates into RMSE of hourly GHI estimates, but because the binary presence/absence of clouds is used with a constant radiation attenuation coefficient (40% of clear sky GHI) it is likely that instantaneous estimates can vary up to that number.

Success of TSI-based forecasting depends on accuracy cloud detection algorithm, and correctness of forecasted 2-dimentional cloud mask. The latter is challenged by the lack of information on 3-dimentional structure and multi-level dynamics of the observed clouds. Perspective and occlusion effects on clouds, cloud deformation and heterogeneity of cloud velocity at different altitudes are the major sources of error in this approach [16]. The approach is most successful in tracking single layer broken clouds moving across the sky without rapid deformation and becomes redundant in clear sky conditions as well as under overcast.

4) *Statistical approaches to short-term forecasts*

Reikard2008 reviewed a number of statistical approaches to solar radiation forecasting, such as log regressions, Autoregressive Integrated Moving Average (ARIMA) models, unobserved components models, transfer functions, neural networks, and hybrid models. Nearly in all tests the best forecasting results were obtained by ARIMA. The autoregressive models capture relationships observed in recorded hourly series of global irradiation between pairs of data points separated in time, e.g. the relationship between the value at one hour and the value at the previous hour; or the relationship between the value at one hour in one day and the value at the same hour in the previous day. While ARIMA generated data may have same statistical properties as the real ones (the mean, variances and cumulative probability distribution) it is not clear if the developed models are transferrable to locations with different clearness index [17]. Because of this location dependence in statistical properties a full knowledge of long term datasets is required for recreating statistical behavior of data.

Linking statistical modeling with real-time data from monitoring site would lead to better accuracy of predictions. While 1 minute ground measurements can be downgraded to coarser time intervals, sub-hourly forecasts can not be produced if observed radiation data is available in coarser resolution. It remains to be seen if application of statistical approaches to day-ahead forecasts can outperform NWP-based approaches in long-term tests.

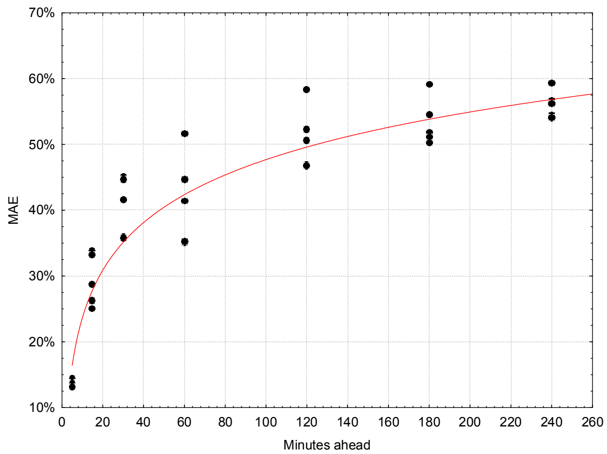

Fig. 3. Plot of average MAEs of 6 different statistical forecasting models applied to 6 locations, based on data from [1].

5) Data limitations in Solar radiation models

Models used to evaluate clear sky solar radiation have limitations, which define the potential maximum accuracy of fully informed forecasts. Solar radiation models produce a small bias under any climatic condition, typically within ±3% for GHI and ±7% for DNI on an annual basis [13]. Evaluations of broadband and spectral radiation models carried elsewhere [18] consider bias of 5% and a RMSE of 8% excellent [20]. Most uncertainties in the models arise from spatial and temporal variability in atmospheric turbidity, which especially in cases of low solar elevations accounts for the largest errors in clear sky irradiance forecasts [18], [19]. Correct estimation of the attenuation of sunlight in atmosphere depends on accuracy of data quality used for estimating scattering and absorption of sunlight, dependent mostly on water vapor, ozone, and aerosols (not accounting for clouds). Good estimates of solar radiation are possible only when turbidity and prcipitable water are well characterized [18].

Because of the high spatial and temporal variability in aerosol optical depth (AOD) there is a strong need for high resolution aerosol information. Incorrect AOD forecasts lead to overestimation of direct radiance and underestimation of diffuse radiance (12% and 14% correspondingly), which results in mean bias of ~2% for GHI [20] when AERONET values are compared to ground measurements. Like aerosols, water vapor is highly variable over space and time. It typically induces uncertainties in the range of 3–5% for DNI or GHI [13]. Total column ozone prediction error is 3 – 4% [21] with previous day ozone as the most informative predictor. Cumulatively, uncertainties in the estimates and forecasts of inputs into solar radiation models have strong effect on solar radiation forecast performance.

6) Ground measurements

Accuracy of ground measurement data defines the potential extent of knowable about the performance of the forecast and of accuracy of persistence-based forecast models. Like any other model, solar radiation models cannot be validated or verified to a level of accuracy greater than that of the measurements. According to Stoffel et al. [22] even in the good measurement regime around noon time, hemispherical field measurement uncertainty is typically two to three times that of direct-beam measurements, or ± 4% to ± 5%, over a year, mainly because of these seasonal uncertainty variations. Measurement uncertainties for pyranometers used to measure GHI in the field range from ± 3.0% for solar zenith angles between 30 degrees and 60 degrees and up to ± 10% for SZA greater than 60 degrees [22].

Inconsistencies in the measured data, resulting from miscalibration, instrument drift, vandalism or lack of cleaning, is a problem in measuring GHI and even greater problem for DNI. Our comparisons of two sets of instruments from the same site shows that dirty instruments (e.g. cleaned once a month and affected by birds) underestimate GHI by up to 5% over a year compared to instruments cleaned every 3-4 days.

7) Location dependence

Forecast accuracy strongly depends on the climatic conditions at the forecast site. Because of the high influence of cloud cover on solar radiation reaching ground, cloud regime strongly defines success of forecast performance. As discussed above, clear sky forecasts at any time scale are relatively accurate, with MBE relating to longer-term variability in water vapor, ozone etc, while cloudy sky forecasts provide more challenge with short-term variability related to types and altitudes of clouds, and their dynamics.

Regionally, difference between sunny and cloudy climates, when tested using same sets of models, is almost double. For Central European stations for example, the relative RMSE ranged from 40% to 60%, while for more sunny Spanish stations relative RMSE values were in the range of 20% to 35% [4]. RMSE in satellite based forecasting is also dependent of average clearness index [12]. In Heliosat model for converting satellite images to solar radiation maps, RMSEs of daily values computed from the measured data of 35 European stations were 20%, 16% and 10% for January, April and July, respectively. Seasonal averages in Iran varied from 22.1% in autumn to 8.4% in the spring (e.g. [23]). Locally, difference in forecast quality between coastal and inland sites demonstrates difficulty in accounting for details microclimate specific to a particular location. In coastal southern California, for example, the marine inversion layer common from May till September, is creating overcast conditions throughout morning until noon. Cloud ceilings of about 1000 m and reduction in average visibilities to 6 miles contribute to MBE of up to 54% on summer mornings [24].

Fig. 4. Marine layer event off the coast of Southern California (GOES visible spectrum image). Note propagation of the low clouds inland.

IV. Recommendations on Industry Standards for Forecasting Performance Evaluation

This short review of capacities of different solar forecasting approaches helps identify a number of common characteristics of forecasts and existing limitations of operational forecast performance at different time scales.

There is an immediate need in using standardized statistical metrics (including data aggregation and time span covered by analysis). Current diversity in the methods of quantifying and expressing forecast uncertainty makes direct comparison of different models challenging. For this reason, fixed levels of RMSE are difficult to use for evaluation of forecast quality. This may help explain why there are no well defined standards in solar forecasting, only suggestions based on arbitrary decisions. The power industry and governing bodies defining forecasting standards for operational or regulatory needs must move from arbitrary (e.g. number of days with "successful" forecast) towards purpose driven accuracy standards, by expressly relating forecast quality to particular management decisions and quantifiable financial gains. Consideration of three broadly defined time scales – operational (minutes to hours), management (hours to days) and short-term planning (days to weeks) may lead to using different metrics at these scales.

For example, for time scales from seconds up to hours ahead the performance metrics of interest should be RMSE because it better addresses the likelihood of extreme values related to ramp-up or ramp-down events. For day-ahead/short term forecast, MAE estimates could be more important because it better relates to unit commitment and energy trading, by integrating absolute difference between expected and observed power production. Medium to long-term forecasts should likely be more concerned with MBE, which is more useful for planning purposes then variability metrics because it relates the best to the assessment of total forecasted power generation capacity.

There is a further need in dialogue between the power industry, forecast service providers and academia on forecast evaluation metrics, time horizons and granularity, and applications of different time scale forecasts to the operational needs of industry. Through this dialogue, an industry consensus could be developed in regard to operational requirements, resulting time horizons and granularity, and applications. Accordingly, appropriate forecast evaluation metrics can be assigned to most closely relate to core operational objectives. As a result of industry-wide collaboration, the standards can better relate to the power industry's decision making process as a part of cost-benefit analysis applicable to different temporal scales.

There is a need in flexibility and account for location specificity in defining the desired realistic forecast accuracy. While relatively high accuracies are achievable in sunny climates, the same numbers are not realistic in cloudy locations. Because of this strong location dependence, the currently published best values (e.g. RMSE) can not be used directly for determining a standard, however, should serve as guidelines. We suggest a discount value of 50% for accuracy claims when a forecast model evaluated in a relatively sunny location is applied to another, relatively cloudy location.

As discussed above, currently there are several levels of uncertainty which limit forecast performance: RMSE of 5% defined by the accuracy of groundtruthing instruments (established over long-term measurement time span) is at the lowest limit of meaningful forecast performance. The next uncertainty level is defined by accuracy of good quality clear sky solar radiation models, at approximately 8%. The third level of uncertainty for the best quality forecast is introduced by local climate.

According to published best practices the lowest RMSE (established over long-term analysis time span) for hour ahead forecast is around 15% and for a day ahead – around 20% (normalized by the mean). Using data presented on Fig. 1 and Fig. 2 we have produced a logarithmic fit of best forecasts for 1 hour to 3 days ahead intervals. Assuming that these forecasts were produced by best models at relatively sunny locations, we used them as a guide for defining the upper extent of forecast error by doubling RMSE. We have also assumed that different methodologies are smoothly blended across intraday and day ahead forecast scales. The results are presented on Fig. 5 which can be used as a crude guide for evaluating forecast performance at different locations. We propose that forecasts producing RMSE values falling between the two lines corresponding to mostly clear and mostly cloudy climates are demonstrating good performance within a range of 25 – 75% clear sky days a year.

Forecast service providers and academia, through collaborative efforts, need to provide in depth research as to the limits of best achievable performance in solar radiation forecast, and spearhead new approaches capable of

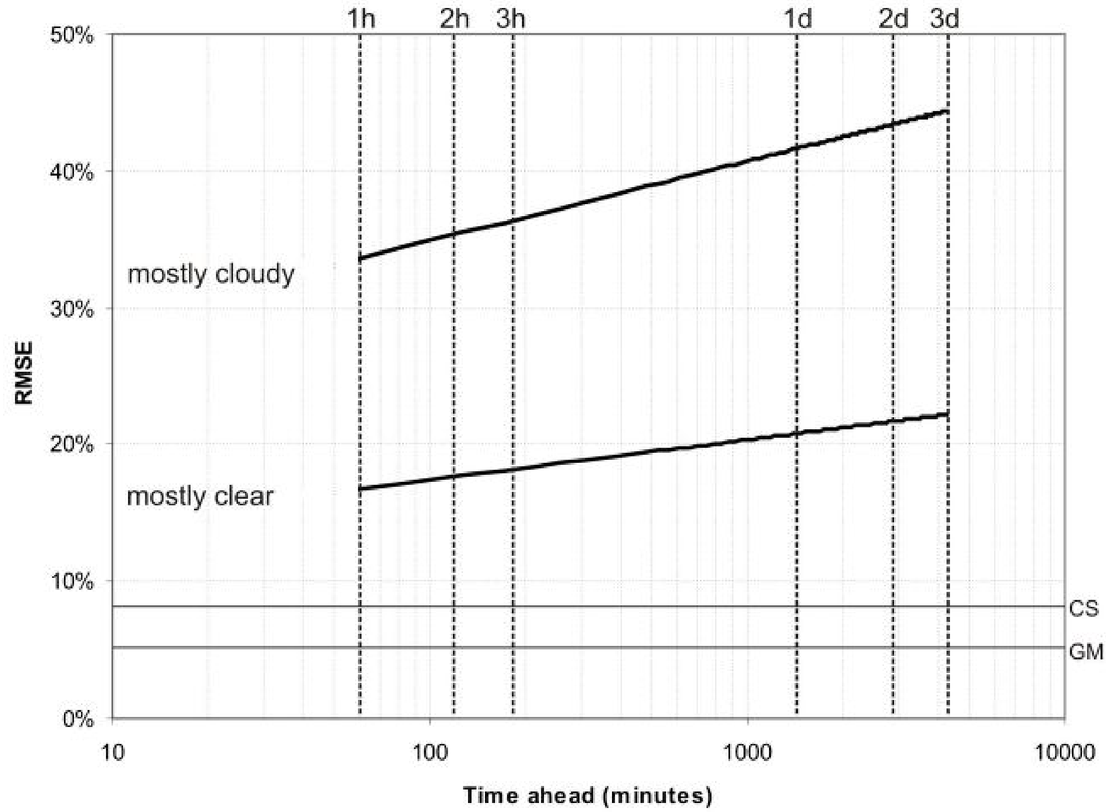

Fig. 5. Conceptual guidelines for evaluation of current and expected forecast model performance. Two lines are shown – the line marked as "mostly clear" indicates logarithmic fit for lowest RMSE values normalized by mean from [9] and [4] which likely result from combination of best model and best climatology. The line above indicates suggested expected RMSE of good quality forecasts in mostly cloudy climates (double RMSE). Horizontal line marked as SC is a RMSE level for good quality clear sky radiation model, and line marked as GM stands for uncertainty of ground measurements.

breaching the existing limitations in forecasting clouds and turbidity-related variables.

While it seems that solar radiation forecasts can be conceptualized as a simple set of technologies, the variability in microclimates, locations, power system designs and user needs turns forecasting into a research project in each particular case.

Competitive forecast trials as an integral part of forecast vendor selection is recommended as industry best practice at the current stage of solar forecast technology maturity. The optimal way to acquire the best forecasting service is not to rely on published (and often forward-looking) statements on forecast accuracy but to allow several forecasting companies to provide forecasting service for the same facility at the project development stage, carry performance analysis for an extended period of time, preferably covering the full range of weather conditions characteristic to the site, including tests of ground measurement equipment already installed for validating forecast performance, and reporting on a same set of accuracy measures applied to same analysis intervals.

V. Acknowledgment

The authors gratefully acknowledge the contributions of Marlene Moore and Tony Daye from Green Power Labs Inc (GPLI) for their help with this manuscript, insight, support and accessing industry needs in solar forecasting. We thank Jack Peterson (SCE), Jim Blatchford (CAISO) and Dallas Cormier (SDG&E) for providing advice and feedback on our forecasting work. We also thank Trevor March (Green Power Labs Australia) and GPLI staff – Jim Fletcher, Shawn Huang, Chad Mahoney, Andrey Kostylev, and Chris Carter for their persistent work on developing, implementing and analyzing performance of solar forecasting approaches.

VI. References

[1] G. Reikard, "Predicting solar radiation at high resolutions: A comparison of time series forecasts," *Solar Energy, vol.* 83, pp. 342–349, 2009

[2] T. Muneer, M. S. Gul, "Evaluation of sunshine and cloud cover based models for generating solar radiation data," *Energy Conversion & Management* , vol. 41, pp. 461-482, 2000

[3] A. Hammer, D. Heinemann, C. Hoyer, R. Kuhlemann, E. Lorenz, R. Muller, H. G. Beyer, “Solar energy assessment using remote sensing technologies,” *Remote Sensing of Environment,* vol. 86, pp. 423–432, 2003

[4] E. Lorenz, J. Remund, S. C. Müller, W. Traunmüller, G. Steinmaurer, D. Pozo, J. A. Ruiz-Arias, V. L. Fanego, L. Ramirez, M. G. Romeo, C. Kurz, L. M. Pomares, C. G. Guerrero, “Benchmarking of Different Approaches to Forecast Solar Irradiance,” *24th European Photovoltaic Solar Energy Conference*, Hamburg, Germany, pp. 21-25, September 2009

[5] A. Hammer, D. Heinemann, C, Hoyer, E. Lorenz, “Satellite Based Short-Term Forecasting of Solar Irradiance - Comparison of Methods and Error Analysis”, in *Proceedings of the EUMETSAT Meteorological Satellite Data Users' Conference*, October 2001

[6] P. Mathiesen, J. Kleissl, “Evaluation of numerical weather prediction for intra-day solar forecasting in the CONUS,” *Solar Energy,* vol. 85, pp. 967–977, 2011

[7] R. Perez, K. Moore, S. Wilcox, D. Renné, and A. Zelenka, “Forecasting Solar Radiation – Preliminary Evaluation of an Approach Based upon the national Forecast Data Base,” *Solar Energy,* vol. 81, 6, pp. 809-812, 2007

[8] M. Lave, J. Kleissl, “Solar variability of four sites across the state of Colorado,” *Renewable Energy,* vol. 35, pp. 2867–2873, 2010

[9] R. Perez, S. Kivalov, J. Schlemmer, C. Hemker, and T. Hoff, “Short-term irradiance variability: Station pair correlation as a function of distance,” in *American Solar Energy Society – Proc. ASES Annual Conference*, Raleigh, NC, 2011

[10] R. Perez, P. Ineichen, K. Moore, M. Kmiecik, C. Chain, R. George, F. Vignola, “A new operational satellite-to-irradiance model,” *Solar Energy,* vol. 73, pp. 307–317, 2002

[11] C. A. Gueymard, S. M. Wilcox, “Assessment of Spatial and Temporal Variability in the U.S. Solar Resource from Radiometric Measurements and Predictions from Models Using Ground-Based or Satellite Data,” *Solar Energy,* vol. 85, pp. 1068-1084, May 2011

[12] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker Jr., D. Renné, T. E. Hoff, “Validation of short and medium term operational solar radiation forecasts in the US,” *Solar Energy* vol. 84, pp. 2161-2172, December 2010.

[13] T. Cebecauer, M. Suri, C. Gueymard, “Uncertainty sources in satellite-derived Direct Normal Irradiance: How can prediction accuracy be improved globally?” in *Proceedings of the SolarPACES Conference*, Granada, Spain, 20-23 Sept 2011, 8 pp.

[14] I. A. C. Martins, P. M. Ferreira, A. E. Ruano, “Estimation and Prediction of Cloudiness from Ground-Based All-Sky Hemispherical Digital Images,” in *Proceedings of CONTROLO'2010, the 9th Portuguese Conference on Automatic Control*, September 2010.

[15] P. M. Ferreira, I. A. C. Martins, A. E. Ruano, “Cloud and Clear Sky Pixel Classification in Ground-Based All-Sky Hemispherical Digital Images,” in *Proceedings of CMTEE'2010, the IFAC Conference on Control Methodologies and Technology for Energy Efficiency*, vol. 1, March 2010.

[16] C. W. Chow, B. Urquhart, M. Lave, A. Dominguez, J. Kleissl, J. Shields, B. Washom, “Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed,” *Solar Energy,* vol. 85, pp. 2881–2893, 2011.

[17] L. L. Mora-Lopez, M. Sidrach-De-Cardona, “Multiplicative ARMA Models to Generate Hourly Series of Global Irradiation.” *Solar Energy,* vol. 63, pp. 283–291, 1998.

[18] C. A. Gueymard, “Direct solar transmittance and irradiance predictions with broadband models. Part II: validation with high-quality measurements,” *Solar Energy,* vol. 74, pp. 381–395, 2003.

[19] C. A. Gueymard, “Importance of atmospheric turbidity and associated uncertainties in solar radiation and luminous efficacy modeling,” *Energy,* vol. 30, pp. 1603–1621, 2005.

[20] H. Breitkreuz, M. Schroedter-Homscheidt, T. Holzer-Popp, “A case study to prepare for the utilization of aerosol forecasts in solar energy industries,” *Solar Energy,* vol. 81, pp. 1377–1385, 2007.

[21] D. Spankuch and E. Schulz, “On Short-Term Total-Column-Ozone-Forecast Errors,” *Atmospheric Environment,* vol. 31, pp. 117-120, 1997.

[22] T. Stoffel, D. Renné, D. Myers, S. Wilcox, M. Sengupta, R. George, C. Turchi, “Concentrating Solar Power. Best Practices Handbook for the Collection and Use of Solar Resource Data,” *Technical ReportNREL/TP-550-47465*, September 2010.

[23] I. Moradi, R. Mueller, B. Alijani, G. A. Kamali, “Evaluation of the Heliosat-II method using daily irradiation data for four stations in Iran,” *Solar Energy,* vol. 83, pp. 150–156, 2009.

[24] A. Nottrott, J. Kleissl, “Validation of the NSRDB–SUNY global horizontal irradiance in California,” *Solar Energy,* vol. 84, pp. 1816–1827, 2010

[25] M. Lave, J. Kleissl, E. Arias-Castro, “High-frequency irradiance fluctuations and geographic smoothing,” *Solar Energy,* Available http://www.sciencedirect.com, 2011